Langley DAAC User Working Group Meeting Executive Summary

September 11-12, 1997

Compiled by Lin Chambers

The meeting began with a report on the status of open action items. One action item on calibration was closed, as was the action on V0 to V1 data migration. While the latter issue is not resolved, the timeline is so unclear that we will simply reopen this issue when it becomes appropriate.

The Langley TRMM Information System (LaTIS)

Richard McGinnis, DAAC Assistant Manager, gave an overview and status report on this system. The future of EOSDIS is uncertain, and LaTIS is part of this DAAC's strategy to meet that uncertain future. As reported before, LaTIS is being developed using the Rapid Development Method. We are approaching the end of the first cycle of development. The requirements and definition phase for the next cycle will be starting in the next few months. UWG members are invited to be as involved as they like in this process.

The LaTIS schedule is basically unchanged from the last meeting. Mission Sim's 2 and 3 were successfully supported, and a TRMM 30 day test is now in progress (and being extended to 45 days). The system is performing well. Science Software Integration and Test (SSI&T) has been successful for all CERES sub-systems delivered to date. Current efforts are focussed on making the system somewhat more automatic (less labor intensive).

The system was designed to be as similar as possible to ECS hardware and software choices, and McGinnis feels that this was not a bad decision. The scope of the design has expanded somewhat to address issues of schema and "lineage" (basically version control and processing history), and to make it easier to operate and maintain. It has incorporated technology from COTS packages, and is using "Big Sur" (a database approach developed at UC Berkeley). "Big Sur" was developed specifically to support Earth Science data, and has a well refined schema and also addresses the version control problem. Some problems were encountered in bringing this code in, but they appear to have been solved. A subscription service developed at JPL for the Mars Pathfinder mission is being investigated to support standing orders more automatically. Data distribution is being automated in this system. Currently they plan to support 4 & 8 mm tapes, DLTs and CD-ROMs. Any UWG members who would like other media are asked to speak up.

At launch LaTIS will support automated ingest and archive, automated PGE processing, web access to data products, media delivery of products, and production request support. It will offer enhanced subsetting and spatial selection, automated media preparation, and possibly the subscription service.

Problems encountered during development were summarized. Probably the biggest has been the Fortran compiler, which has not performed very well. A new version coming out soon may solve this issue. Software ports to the new SGI Operating System took longer than expected. SGI deliveries were slow and the system was initially very unstable. Informix was slow to deliver the database software; and the AMASS software interfacing to the large storage devices did not perform as advertised.

Successes from the development include learning to use COTS in a heterogeneous environment, supporting all TRMM tests to date (even some that were not in the initial plan), and staying on schedule.

Lessons learned (again) are that vendors lie, and database vendors are in addition very greedy. Taking research prototypes to production can be tricky; and lots of contingency plans are necessary and must change often.

As this development proceeds, some hardware elements of the V0 system are becoming unsupportable (they're already 4 years old!), so the DAAC anticipates that all V0 data will end up on LaTIS.

Jill Travers gave an overview of LaTIS SSI&T. About 1/3 of the PGE's for TRMM have been delivered and tested to date, using 3 different compilers so far. It's planned to move everything to SGI compiler 7.2 soon.

We then moved to the DAAC building where a demo of the CODINE job scheduler software was given. We then toured the DAAC hardware room upstairs. An impressive number of Origin 2000 CPU's are in place - and actually getting some use - for LaTIS. A Storage Tek is also in place. The tape writers and DLT writer are in place, while the CDROM writer is expected. A large bank of working storage is also in place. The ECS Testbed hardware is also in place.

Data Access

Bob Seals gave an overview of the state of data access at the Langley DAAC. The current approach is to provide a variety of access methods, so that all possible users can be served. During this discussion it was pointed out the the DAAC may have reached a transition point. Data access issues are being appropriately addressed, and now increased emphasis needs to be given to the challenge of making the data more generally usable (i.e., problems with HDF or any particular format; platform compatibility issues). Ideas for doing this were explored, but there is no simple fix to the problem. As was pointed out, this issue has been around ever since there has been more than one computer. In discussion later in the meeting, we did settle on one action item:

Action: ESDIS to open listserv to discuss problems reading data. If such a capability already exists (HDFHELP?) then they should advertise it better. Archival of past messages should be included.

A number of options to address this problem were also identified:

1. keep track of who is having what kinds or problems.
2. translator/subsetter software
3. translate on the fly as data are ordered
4. data providers to provide software to read their data (this needs attention)
5. something besides HDF; or if not, at least some user-friendly documentation (i.e., a step-by-step guide to installing HDF libraries and running tools)
6. offer custom translation services (probably not practical)

A demo was performed of the directly downloadable data, which is accessible from a simple web page at the DAAC. It was the determination of the UWG that all projects should be listed on such a page; even if the data can only be ordered (this decision was borne out the following week in discussion with someone from the ESDIS project: many people search only using web search engines, and the DAAC's data need to be findable that way).

A demo of the prototype web search interface was also performed. This interface is currently in development, and the group discussed the pros and cons of early release to the user community of a prototype that doesn't always work. It was determined that, unless it **never** works, it is worth getting out to the user community as soon as possible for comments and feedback. There was also discussion of the ECS web prototype (JEST) and whether the effort put into it was justified by the number of users. It was too late in the development, however, to make any change here.

Strong Endorsement: The UWG registers a strong endorsement of the current web approach being developed by the DAAC. Multiple approaches are important and web search engines must be able to find all the datasets held at the DAAC. The low end user must be kept in mind (reasonable bandwidth and technology requirements).

ECS/EOSDIS Status and Plans

Lucy Lee gave a review of the ECS Pre-Release B Testbed Status. She reported that this testbed has established a foundation for a viable operations environment. The configuration change process has been put in place, with 250

trouble tickets written. Of these, 34 are still open; 53 have been promoted (to the next release). This process seems to be working well. A repeatable system upgrade and installation process has been established as a result of recent efforts, and a foundation has been laid for technical support procedures.

Jill Travers gave an overview of Science Software Integration and Test (SSI&T) on the testbed. Two Product Generation Executables (PGEs) have been run through the system for CERES, but this effort is at a lower priority than SSI&T activities on the LaTIS system. One MISR PGE (of three expected) has been tested. For this PGE, 3 of the 4 test cases have run successfully through the PDPS system. The 4th is on hold while other activities take place. There has been an issue with the use of future time in the system, which may be resolved by now. Three MOPITT PGEs have been delivered. All run at the command line, but some changes were required to get them to run through PDPS. No SAGE PGEs have yet been tested. It is still unclear whether SAGE will be given approval to do their processing at the SCF instead of the DAAC. [Note: this approval was given on 9/24/97.]

August Demo

Richard McGinnis attended the August Demo of the ECS system, and gave his overview of the event. The EOSDIS Review Group (ERG) recommended that this demo be enhanced to support a Go/NoGo decision on ECS, while alternative plans were made in case of a NoGo decision. The demo was held on August 28. A meeting of those who saw the demo was to take place at GSFC in late September to provide input to the ERG. A formal decision was expected sometime in October.

The August demo system consisted of 500,000 lines of C++ and 70 COTS packages. Three scenarios using LANDSAT, MODIS, and ASTER demonstrated the functions of the system.

McGinnis reported that the demo was well-orchestrated, and that Hughes had definitely made progress. All the criteria for the system were basically met, and it also dealt automatically with an unexpected error. The system ran slowly, but this was blamed on use of a SUN machine for the dataserver. It will be on an SGI later, and is expected to be 10 times faster. The system design is event-driven and uses a subscription service. The performance measurement plan (this aspect was not part of the demo) was inadequate, as it will measure subsystem performance independently. There is no plan to test end-to-end performance of a heavily loaded system (the typical operational condition). There is also no plan to move to SGI Origin machines (which form the basis of the LaTIS system). While basically successful, this demo did not allow for any determination of performance or robustness of the operational system. It also did not give any indication of the difficulty of operation (and hence the number of operations staff needed at the DAAC). At most two things were run at any one time during the demo.

As a result of the demo, McGinnis summarized his concerns:

- Hardware will be insufficient (due to past budget cuts and current planning. Also, SGI Challenge saturates under much smaller I/O load than SGI Origin).
- Planned performance tests are questionable (will derive rather than measure overall performance).
- Planned DAAC Operation Staff may not be sufficient (no way to tell)
- If federation happens, how will ECS be maintained?
- No decision will be made (ECS will continue with DAAC's supporting alternative, emergency, and other plans in parallel, leaving DAAC with no reserves left).

Post-August Demo Focus

Lucy Lee reported on the focus of ECS activities following the August demo. Test plans are being streamlined to help meet the required schedule: redundant tests are being eliminated while concurrent testing is planned wherever possible. Scenario-based testing is being used to test the right things. Incremental deliveries are being planned now, rather than the original "big bang" view. The emphasis is on fast, reliable turnaround for fixes and enhancements. For operations, integrated status and error logging are being worked on (not the current situation), and a consistent user environment setup is in the works.

AM-1 Operations Readiness

Lucy Lee also reported on the readiness for ECS AM-1 operations. Issues outstanding are testing, training, operations agreements, and the AM-1 Operations Readiness Review (ORR). Test re-planning was mentioned in the previous section. There are some issues about where and when training for DAAC staff will take place, and how this can be supported logistically, given the other operational responsibilities the Langley DAAC has. Operations agreements are currently being developed in concert with the instrument teams (ITs). The DAAC has had active participation in the DAAC Operations Working Group (OWG) to prepare for AM-1, and has also established a joint CSC/ECS operations support team.

The AM-1 ORR is to be held in the May/June 1998 time frame, and will determine the DAAC's readiness for AM-1 science operations and continuing V0 and TRMM operations. ITs and the ESDIS project will have to concur on readiness.

ITs are also responsible for delivering science software on schedule, and **for following the published COTS baseline** and software standards. The DAAC does not have the resources to continually upgrade to new COTS software releases, or to support multiple versions of software releases for individual instrument teams. (It was pointed out by an IT member, however, that there have been instances where the baseline has had to be adjusted, due to problems with compilers. Decisions about what to support have to be made on a case-by-case basis where benefits, resources, etc are weighed.)

Jim Frenzer reported that the TRMM ORR is to be held at the end of October, 1997 and will determine the DAAC's readiness for TRMM science operations and continuing V0 operations.

The DAAC and the ITs have also been participating in "Emergency Planning", in the event that ECS should slip more.

Possible Future Datasets

A vote was taken on the nomination of the ER-2 MAS data from ARESE. Hearing no objections, this dataset was approved for inclusion at the DAAC. Paul Bailey is still working on the possibility of archiving some NCAR Tropospheric chemistry datasets here; there has been no committment of interest yet. There were no new dataset nominations at this time.

NSSDC Archive

Bob Seals distributed some additional information on the order history for these datasets (see complete list or sorted by more recent requests). The ensuing discussion revealed little interest in any of these datasets. Some UWG members specifically recalled that certain datasets were not of much use. In particular, those that cover only a short time period are probably not much use, and the coarse resolution of many of the early products limits their use now. Some were, however, singled out for discussion. It was suggested, for example, that the LIMS RAT data might be put on a CD-ROM and kept (at the GSFC DAAC?). Some of these datasets may also already be archived in some form at NCAR (Roy Jenne has been contacted to inquire about this). It was suggested that Dr. W. Smith, formerly of U-WI, also be polled on these datasets (a message has been sent). Subsequent e-mail discussion with the rest of the UWG also noted: NIMBUS-7 (most recent requests, long duration datasets); early MAPS data (should check with PI to discover the status of this).

DAAC Status

Dataset Activities

John Olson summarized dataset activities at the DAAC since the last UWG meeting. 204 of 205 datasets now have on-line guides. The next priority will be source, sensor and project guides. DIFs (Directory Interchange Format) are

different enough from guides that they must be prepared separately. 168 DIFs have been completed and 40 are in progress for the 205 datasets on-line. Seven supplementary information templates (SITs, which provide background information) have also been completed, with 10 more in progress. These do not correspond one-to-one to the on-line datasets.

In the 6 month period of March-September, 1997, 155,519 MB of data were added to the DAAC in 32 different datasets. A few of these datasets (GTE) are available only through the web page; and not through the V0 system. As of September, 1997, total holdings at the DAAC were 848,075.54 MB in 19,279 granules. A SAGE II CD-ROM was also completed and released in March, 1997.

Usage Statistics

Charts summarizing usage of the DAAC since inception were also provided. There is generally an upward trend in both the number of users and where they come from; as well as in the volume of data distributed. Summary information by quarter for 1996 and 1997 illustrates the point that media distribution is a major factor. It was pointed out that the "bimodal" pattern of usage still continues: a few users order large amounts of data, while many (~90%) of users order small amounts.

Outreach Update

The USWG CD-ROM was not discussed. Its development is progressing.

Newsletter

A draft of the fall newsletter (Vol. II, No. 2) was distributed. Topics in the newsletter so far have included: educational outreach, datasets, CD-ROMS, a DAAC Focus story (i.e., summer students at the DAAC), and upcoming events/conferences. Suggestions for future topics are welcome. The current draft contains an article about job openings at the DAAC, which was under review for legality. Some of the difficulties the DAAC has in recruiting qualified candidates (with knowledge of atmospheric science, etc...) were discussed. The UWG was invited to help identify potential recruits. Distribution of this newsletter is still planned on a quarterly basis, via surface mail. The mailing list currently consists of anyone who has ordered data between January 1996 and January 1997, EOSDIS DAAC User Services Offices, and UWG members. It was suggested that it should also be sent to all NCAR members, and to super computer user services, as well as being published on the website.

Trading Cards

Two K-12 teachers worked at the DAAC full-time this summer. Their principal product was a set of trading cards, with associated lesson plans, on the Earth Radiation Budget. Future card sets are in development covering the other discipline areas for the DAAC. The trading cards themselves are targeted as a stand-alone product for a grade 5-8 audience. Each card includes a URL which will lead the interested student to further information. Lesson plans for teachers are also available, relating ERBE to science, technology, language arts, and math. This trading card set is currently in final review in the NASA Langley Office of Education, and it is hoped to have the initial printing before the end of the calendar year.

S'COOL

An update was given on the CERES S'COOL Project. This project has taken off since the last UWG meeting, with 8 more schools participating in Phase 2 testing in April. Phase 3 over the summer fizzled, due to the amount of time required to get things going with teachers in the southern hemisphere. However, work continued during that time on automating S'COOL processing to accommodate more participants, and an effort to translate the website into useful languages was begun. Recruitment for Phase 4 in October proceeded well (as of this writing, 31 schools - 17 in US, 14 in Europe - will be making observations the week of Oct. 20, 1997). Some print materials have been drafted, to reach

the less computer- savvy teachers. Sample teacher brochures were handed out. Final versions of the brochure and a color classroom poster will be printed following feedback from Phase 4, and will be available for anyone who is interested. As we approach the operational phase of CERES, and hence of S'COOL, articles are being submitted to various teacher magazines and journals to make them aware of the project.

DAAC Work Plan

Jim Frenzer discussed the DAAC work plan for FY 1998. The work plan includes science support objectives, past year accomplishments, and plans for the coming year. It also covers out-year projections, and includes milestone and budget information, and staffing projections. Feedback from last year's plan is being used in preparing the current plan: systems engineering focus was excellent, tables were used to summarize information; but it was much too long with too many details and needed more emphasis on services. The LaRC DAAC was the only one to submit its plan on time last year (it will not have that distinction this year, since it's already late).

For the 1998 plan, some of the systems information will be repeated, but detailed information on SSI&T and Mission Support will be removed. A 2-page milestone template from ESDIS was distributed and will be used as the basis for setting schedules. However it is recognized that these milestones are a moving target. The section on Science Support Objectives will be redone to emphasize the focus on users.

Items to be included in the plan are: TRMM operations (24 hours a day, 7 days a week, but with efforts to minimize off-shift personnel); finish 1st phase of LaTIS; prepare for AM-1 and Meteor operations (using yet-unseen ECS system, with open issues regarding timing and location of testing and training); and continue previous work (user services, outreach, deal with existing and future datasets).

To handle all this work, the DAAC staff (as of 9/12/97) has 66 members (5 NASA, 55 CSC contractors, 6 Hughes contractors). Turnover has been fairly low: since June 1996, 30 people have been hired and 8 have left (including 3 of the new hires). There are several positions open on the CSC staff, and this led again to a discussion of how the positions are and should be advertised, and how the UWG may be of help in the process. Projections call for about 25 more people in the next year.

Once a few more things are complete, it is planned to distribute the work plan to the UWG for comment.

NRC Recertification and ESIP CANs

The NRC study of EOS recommended a few years ago that EOSDIS become instead an Earth Science Information System, composed of a federation of partners (ESIPs). NASA developed a plan to implement this recommendation, involving three types of ESIPs. Cooperative Agreement Notices for Types 2 (developmental and flexible) and 3 (outside of NASA GCRP) were released on May 8, 1997, with proposals due July 14. A decision on selection is due in October, 1997. The existing DAACs are to be recertified by the NRC and become the baseline federation to deal with TRMM, AM-1, and PM-1. All work after those flights would be competed.

Recertification of the DAAC's was to be done by the NRC, to ensure that the DAACs are doing a good job (looking at usage and trends, staff expertise, user support, user satisfaction, outreach efforts and effectiveness, and services/technology added by the DAAC). The review of LaRC will take place Nov. 17-18, 1997, and criteria for recertification are currently being negotiated.

UWG Status

Only one formal action item, on ESDIS, was assigned during this meeting. A general endorsement of the DAAC's approach to data access was given, with a general charge to now start giving more emphasis to the useability of the data once it has been accessed. In addition, the UWG undertook to provide further input on the NSSDC Archive data over the next few weeks.

UWG Charter

A discussion of the UWG charter ended the meeting. This document was written several years ago, and due for revisiting. Changes are suggested below, and will be made before the next meeting unless there are objections from any UWG members.

Some revisions to the objectives were suggested (revisions are **bold**):

1. assist in defining **how the Langley DAAC will meet** the science goals and questions which DAAC activities will support.
2. provide **customer/user/___** ...
3. provide **customer/user/___** ...
4. facilitate science coordination with the general user community, the EOSDIS **Data** Panel, the EOSDIS Program and Project Scientists, **and interested members of other communities**

The term used in items 2, 3 and 4 should include all types of users (not just science or research) but without bringing in unintended meanings. User is the best term that was brought up so far. Other suggestions are welcome.

No changes to the responsibilities were identified, since the list given is not all-inclusive.

It was noted that the membership section calls for 2 year terms with a possible one year extension. Some of the current UWG members have been around longer than that (I think), and should be giving consideration to the nomination of a suitable replacement. The whole issue of how rotating membership is to be accomplished might deserve some discussion by the group. For starters, I'd like to add a "Member since..." line to each entry of the membership list. **UWG members please get back to me with a date for you (or your best guess).**

Changes were suggested to the rest of the section on membership (again **bold**): "... Members are selected to be representative of both the existing user community and the expected EOS community **[cut]**. EOS investigations are represented by members from at least one instrument or Facility team and one Interdisciplinary Scientist team. The DAAC Program Scientist from NASA Headquarters, **the DAAC manager**, and an **ESDIS Project Representative** serve as ex-officio members of the UWG, and ad hoc members may be appointed by the UWG as specific needs warrant. The UWG is co-chaired by a non-DAAC member and the DAAC Project Scientist."

There was some discussion about this last point. If anyone would like to volunteer to serve as co-chair, please contact me. This position would mainly involve helping to set the agenda for meetings, and helping find suitable new members. If no one volunteers, I may appoint someone.

It was also suggested that the following be added:
Substitutes should be sent if a member cannot attend a meeting.

Regarding meetings and interfaces, it was pointed out that the co-chairs are to schedule meetings in a timely fashion to address issues. I would like to emphasize that if anyone thinks a meeting is appropriate at a given time, they can get in touch with me to start planning one. We also noted the last point: "The UWG will designate one member to serve as an ex-officio member of the EOSDIS **Data** Panel." At present, Bruce Barkstrom is a member of this panel, but it is not clear that his position is an ex-officio one in which he represents the DAAC UWG. This issue may require some further discussions with ESDIS. It was also brought out that our participation in the NRC recertification would be welcome and perhaps required in some form.

One change was suggested to the section on Funding and Administration: "...these funds will be **controlled** by the Project Scientist."

Following the discussion of the charter, some brainstorming took place about what should be done at the meetings, and how to make them more attractive to the members. Some ideas: have a presentation during each meeting from one of the members (how I use DAAC data, upcoming field programs, etc...); have an optional day for Instrument Team

reports. For the next meeting in particular, reports on TRMM, CERES and S'COOL were suggested; possibly with video of the launch and even with some teachers/students invited to report on how they view events through the S'COOL Project. Some discussion also took place about whether the UWG as a group should get involved in sponsoring outreach products (so far, the answer is no; but we may continue to discuss this).

Dates for the next meeting have not been set yet. The experiment with doing this last time was rather disastrous, so we will wait until an appropriate time and still try to have a month or two to plan.

UWG Home Page
